Ethnic and class residential segregation: exploring their intersection – a multilevel analysis of ancestry and occupational class in Sydney

Kelvyn Jones, Ron Johnston, James Forrest, Chris Charlton and David Manley

Abstract

Most studies of ethnic residential segregation recognise that occupational class is an important influence on the intensity of segregation of members of different ethnic groups, but are unable to explore variations in that intensity because of the lack of relevant data. Australian census data allow the class structure of different ancestry groups to be identified in small areas within cities. Such data for seventeen ancestry groups in Sydney are used here to explore variations in segregation levels between classes within ancestry groups at three separate scales. To do this, a major extension to a recently-developed methodology for exploring multi-scale segregation patterns is introduced. The results show that for some groups class is more important than ancestry as influences on segregation levels whereas for others there is relatively little class segregation.

Keywords

Segregation, ethnicity and class intersection, multilevel modelling, Sydney, scale

Introduction

In a recent challenge to students of urban residential segregation patterns, Kapoor (2013) has pointed out that most of their empirical studies treat patterns of ethnic and occupational/class segregation separately, ignoring any intersections between the two. They fail to explore whether there is class segregation within ethnic groups, for example, or whether much of the segregation of ethnic groups can be accounted for by their different class structures. This major lacuna in the literature reflects two problems. First is the paucity of data with which the nature of such intersections can be explored: as Reardon et al. (2017, 36-37) note, 'Measures of measuring multi-dimensional patterns of segregation, such as the joint distribution of race and income among neighborhoods … are less well developed' than single-dimensional measures, and those generally deployed 'do not provide a clear description of the joint distribution of race and income across neighborhoods'. Secondly, even if sufficient data were available, the methodological toolkit traditionally used to explore such data is not capable of providing interpretable evaluations of the relative importance of ethnicity and class within any observed multidimensional patterns.

This paper introduces a statistical methodology which directly addresses that issue. Deploying Australian census data that are made available in a form that allows such explorations, it introduces a novel procedure – building on recent innovations in the multi-scale measurement of segregation – that allows a decomposition of segregation levels by ethnicity and occupation; using it their separate and joint contributions to those levels can be discerned.

The likely intersections of occupational and ethnic segregation in urban residential mosaics are generally appreciated in many studies but like Kapoor – who used British census data to investigate unemployment levels for different ethnic groups according to their degree of concentration in the neighbourhoods studied – they have had to use indirect routes to establish the relative importance of ethnic group and occupational class membership as determinants of segregation levels. (See, for example, Owen, 1995; Harris et al., 2016.) In the United States the availability of income, educational and occupational data within the main ethnic groups has allowed analyses of the relative segregation of those subgroups, showing that whereas race is the main determinant of segregation patterns for African-Americans relative to the distribution of Non-Hispanic Whites, class is a more important influence on segregation for both Hispanics and Asians (Iceland and Wilkes, 2006), though again, as Reardon et al. (2017) note, the measurement of those differences is indirect. (See also Iceland et al., 2005; Massey et al., 2009; Intrator et al., 2016; Reardon et al, 2017.)

The Australian census TableBuilder facility allows data to be extracted on individuals categorised by both their ancestry – the census does not collect data on ethnicity – and their occupation, that can be aggregated at a variety of spatial scales.[1] These data have been analysed to identify segregation patterns at multiple scales using an innovative multi-level modelling procedure (described below) for a wide range of ancestry groups and also for generations within those groups (Johnston et al., 2016, 2017). The analyses reported here use tables from the 2011 census that cross-classify ancestry by occupation for all economically-active and employed adults aged 20-64 (i.e. excluding students, the unemployed and those not in the labour force) for sixteen of the largest ancestry groups in the Sydney urban region (Table 1). Ancestry is determined by responses to a question 'What is your ancestry: provide up to two ancestries only?': we have also included those who claim Australian ancestry only as a comparator group. Occupations are grouped into four categories: salariat – professional and managerial; routine non-manual; skilled manual; and semi- and unskilled manual. As reflects a 21st century economy, the majority of respondents fall into the first two (white-collar) occupational classes, with relatively few in the blue-collar occupations in several of the ancestry groups.

Previous analyses of segregation patterns in Australian and other cities have explored the degree to which ancestral groups are concentrated in particular parts of an urban area at a variety of spatial scales (Manley et al., 2015; Johnston et al., 2016, 2017). All groups in the places analysed are spatially concentrated at a number of scales: in particular there is segregation at a macro-scale of intra-urban regions and, within those regions, further concentration at both meso- and especially micro-scales. Groups cluster in particular segments of the urban residential fabric and, within those macro-regions, in particular districts, suburbs, and neighbourhoods. The Sydney data used here have been aggregated into a hierarchically-structured three-level set of areas (with their mean number of adult residents in the four occupational categories):

- Regions (sub-metropolitan labour market areas) – of which there are 43 (41,490);
- Districts (community areas seen as interacting socially and economically) – 223 (8,000); and
- Suburbs (social areas as designated by the NSW Geographical Names Board) – 618 (2,887).

(Data can also be aggregated to a further micro-scale of over 9,000 neighbourhoods, but the small number of respondents in all four of the occupational classes for many of the ancestry groups makes these unsuitable for analysis in this work.)

The multilevel modelling strategy employed here explores the parameters of multidimensional segregation in an overall statistical framework which allows an inferential approach to identify multiple sources of segregation at multiple scales while taking account of stochastic variation that necessarily accompanies small counts. This is achieved by classifying individuals into a large multi-way table of counts that uses the finest geographical division of areas (suburbs in this case), with those individuals differentiated by ancestry and by class. With these counts as the response variable a number of models of different complexity is fitted to assess the nature of segregation – how intense is it at each of the geographical scales for each type (defined by ancestry and occupational class)? For each cell in the multi-way contingency table at the smallest scale – in this case seventeen ancestry groups by four occupational classes by 618 suburbs – an expected number of individuals is derived from the distribution of the total population: thus if, say, one per cent of Sydney's population lives in the area, one per cent of those in occupational group *y* within ancestral group *x* should reside there. The ratio of the observed to expected number is derived and the distribution of those ratios is modelled according to the strategy set out in Jones et al. (2015).

There are two main sets of output from this approach. The first is a Median Rate Ratio (MRR) value for each ancestry by class group. MRRs can be interpreted as follows. Take the set of 43 regions, for each of which we have a modelled observed:expected ratio for the number of members of the salariat in the Irish ancestral group. Select two of those regions at random and calculate the ratio between the larger and smaller of those two values and store it. Repeat the procedure many times, and thus derive a frequency distribution of ratios. The median value is the MRR. It can be interpreted as the average difference between two regions selected at random in the concentration of group members there – thus if the MRR is 2.0 on average one region will have twice as many members of the group present compared to the expected number, relative to another region. The MRR is thus an index of segregation – the higher its value the greater the segregation at that scale (i.e. concentration into certain regions) – and it has associated credible intervals (CIs, which are asymmetric) that can be deployed in comparing two MRRs to see if they differ significantly from each other. MRRs are also computed for each of the other scales, and provide measures of the intensity of segregation there, net of segregation at all higher scales.

Previous studies have presented MRRs for individual ancestral groups in Sydney at various scales (Johnston et al., 2016) and also for different generations within those groups (Johnston et al., 2017). The present paper presents the MRRs for each of the four occupational classes within the seventeen ancestral groups being analysed. It then outlines a further extension of the method that allows an evaluation of the relative importance of class and ancestry as an influence on the intensity of segregation for each ancestral group, thereby providing a formal, statistically robust, statement filling the lacuna on segregation studies identified by Kapoor (2013).

The second set of outputs from the multilevel modelling procedure is a matrix of correlations that can be used to describe formally the degree to which members of each ancestry-by-class group share the same (regional, district or suburban) spaces. These are reported in the final section of the paper.

Segregation intensity

The expectation from the literatures on class and ethnic segregation is that there should be class segregation within each ancestry group, with the greatest intensity of segregation being for those in the lower status occupational groups, reflecting their more limited choice within the housing market relative to their higher status contemporaries. (On that large and extensive literature see, for example, Darroch and Marston, 1971, Massey 1981, Massey et al 2009, and Butler and Hamnett, 2011.) There may, however, be variations across the ancestry groups: those longer established in the city and culturally less distinct from the majority population (in this case, English-speaking) should be less segregated, in each class, than the more recent, culturally more distinct, ancestral groups. For some groups, segregation by occupational class should be the dominant feature of their distribution within the residential mosaic; for others, there may be segregation by both class and ancestry; and for a third group there may be segregation by ancestry alone, with members of all classes concentrated in the same areas.

Those variations may themselves vary according to the scale of the pattern being investigated. As earlier studies of Sydney and other cities have shown, most ancestral groups tend to be concentrated in the same macro-scale portions of the city. These, in most cases, were the regions in which the groups originally settled and from which relatively few, including later generations, have moved away. If that is the case, then at the regional scale, as analysed here, the intensity of segregation should not vary across the occupational classes. At the meso-(district) and more micro-scales (suburbs) within such regions, however, segregation probably varies between classes: members of each ancestry group's lower status occupational classes should be more segregated into particular districts and suburbs within their favoured regions than their higher status contemporaries.

The MRR values for each occupational class within each ancestry group at each of the three spatial scales, together with the associated CIs, are given in Table 2, and the general patterns are described – showing the MRRs without the CIs – in Figure 1. At the regional scale, there is a clear trend within each occupational class down the columns in Table 1: basically, the Asian ancestry groups are the most segregated and the British, Irish and New Zealand groups are – like those who claim only an Australian ancestry – the least. There are a few outliers from the general trend (made clear in Figure 1) with high levels of segregation for the Dutch, Maltese and Lebanese semi- and unskilled class. The same general pattern occurs at the district scale, again with a small number of outliers for the lowest occupational class (though none as substantial as those at the regional scale). There are more outliers at the smallest, suburban, scale – almost all of them referring to one or both of the two 'blue-collar' occupational classes, many of which are small with mean numbers across the 618 suburbs of less than five (the largest of those outliers are excluded from the relevant diagram in Figure 1).

There are major differences within each ancestral group in the relative levels of segregation for the four occupational classes. At the regional scale, there is little or no trend across the four classes for a number of the groups, with no statistically significant differences between the MRR values. The main exceptions to this general pattern – or lack thereof – are for the six ancestry groups shown in the lowest part of the table: Greek, Maltese, Lebanese, Indian, Chinese and Korean. In all six cases, the MRR values show that the semi-/unskilled workers are statistically significantly more segregated than the salariat (the CIs for the two classes do not overlap), which is not the case for most of the other groups.

The same general pattern occurs at the district and suburb scales, although at the latter scale, as noted above, some of the MRR values are based on small numbers of observations per suburb, indicated by the wide gap between the low and high CI values around the MRR in the table. Thus there are several ancestry groups – including the Australian, British and Irish – with no substantial, let alone statistically significant, difference between the MRRs for the four occupational classes at the district scale. For others – notably the Lebanese, Indian, Chinese and Koreans – those in the lowest status of the four classes are significantly more segregated than either the salariat or the routine non-manual classes; with those ancestry groups there is clearly more segregation of (the smaller number of) blue-collar than white-collar workers.

The intersection of ancestry and class; the extended MRR modelling

In order to establish the relative importance of ancestry and occupation in the patterns of segregation outlined above, in principle it would be possible to analyse all seventeen ancestry groups as well as the four occupational classes simultaneously but that would involve a very large number of parameters.[2] Consequently we decided to model the host Australian population plus each of the remaining sixteen ancestries in separate models. This makes for much less unwieldy estimation and allows us to see the relative importance of ancestry and class for each group. Consequently we estimated a set of models of increasing complexity for each ancestry and the Australians – five models in all for each of the sixteen non-Australian ancestries.

Jones et al. (2015) discuss the general specification of the Poisson multilevel model for estimating the degree of segregation. The distinctive feature of the extension introduced here is that we investigate two aspects of segregation – ancestry and class – simultaneously, fitting a number of simpler and more complex models to the same data to ascertain the appropriate specification. For illustration, we start with the most complex model for two ethnic groups (British and Australian) and the four classes (Salariat, Routine Non Manual, Skilled, and Semi-/Unskilled) for only two geographical scales (Suburbs within Districts). This requires a three-level model:

$$O_{ijk} \sim Poisson(\pi_{ijk})$$

$$Log_e(\pi_{ijk}) = Log_e(E_{ijk}) + \beta_{1jk}SalBrit_{ijk} + \beta_{2jk}RNMBrit_{ijk} + \beta_{3jk}SkillBrit_{ijk} + \beta_{4jk}UnskillBrit_{ijk} + \beta_{5jk}SalAust_{ijk} + \beta_{6jk}RNMAust_{ijk} + \beta_{7jk}SkillAust_{ijk} + \beta_{8jk}UnskillAust_{ijk}$$

$$\begin{aligned} \beta_{1jk} &= \beta_1 + v_{1k} + u_{1jk} \\ \beta_{2jk} &= \beta_2 + v_{2k} + u_{2jk} \\ &\vdots \\ \beta_{7jk} &= \beta_7 + v_{7k} + u_{7jk} \\ \beta_{8jk} &= \beta_8 + v_{8k} + u_{8jk} \end{aligned}$$

$$\begin{bmatrix} v_{1k} \\ v_{2k} \\ \vdots \\ v_{7k} \\ v_{8k} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{v1}^2 & & & & \\ \sigma_{v12} & \sigma_{v2}^2 & & & \\ \vdots & & \ddots & & \\ \sigma_{v17} & \sigma_{v27} & & \sigma_{v7}^2 & \\ \sigma_{v18} & \sigma_{v28} & \cdots & \sigma_{v78} & \sigma_{v8}^2 \end{bmatrix})$$

$$\begin{bmatrix} u_{1jk} \\ u_{2jk} \\ \vdots \\ u_{7jk} \\ u_{8jk} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{u1}^2 & & & & \\ \sigma_{u12} & \sigma_{u}^2 & & & \\ \vdots & & \ddots & & \\ \sigma_{u17} & \sigma_{u27} & & \sigma_{u7}^2 & \\ \sigma_{u18} & \sigma_{u28} & \cdots & \sigma_{u78} & \sigma_{u8}^2 \end{bmatrix}$$

$$Var(O_{ijk}|\pi_{ijk}) = \pi_{ijk}$$

where O_{ijk} is the long stacked vector of the observed count for 'individuals' (see later) *i* in Suburb *j* in District *k*. These observed counts, obtained from the most complex models, are used in all subsequent simpler models. The other observed variable is the expected counts (E_{ijk}) for each ancestry by class group, derived as if their numbers were distributed evenly according to the total population size of the lowest level areas (Suburbs) and the city-wide proportion of each group. This represents the count associated with an even distribution of no segregation. There are eight separately coded dummy (0/1) variables (e.g. $SalBrit_{ijk}$) that identify which count represents which ancestry by class group.

We assume that the counts have a Poisson distribution with a mean rate of π. However, it is the natural log of the underlying rate that is modelled, which is achieved using an offset that is the log of the expected count with a coefficient constrained to 1 (McCullagh and Nelder, 1989). The terms in this fixed part of the multilevel model (Duncan and Jones, 2001) are eight estimated means associated with the dummies in the model; β_1 gives the log average ratio across Sydney for the British Salariat while β_8 is the equivalent for the Semi-/Unskilled Australians. As the sum of the observed counts will be equal to the sum of the expected, the expectation is that these means on the log scale will be zero and therefore will, when exponentiated, give the all-Sydney ratio for the mean area as 1. Differentials are allowed to vary around these averages so that v_{1k} is the log differential for District *k* for the British Salariat. If this value is positive there are more of such people there than expected from an even distribution; if it is negative there are fewer. There is a further set of differentials at the Suburb level so that u_{8jk} is the differential for Semi-/Unskilled Australians for Suburb *jk* from District *k*, (v_{8k}) which is itself a differential from the average (β_8) across the city for this group. In this way, the observed values are separated into an overall average and differentials at each scale in the hierarchy, with the difference at the lower scale an estimated departure from the higher scale unit to which it belongs.

Turning to the random part of the model, the differentials (v_{1k} etc.) at each of the two higher levels are assumed to come for a multivariate Normal distribution so that the variance σ_{v1}^2 summarizes the differentials for the British Salariat at the District level. These variances are the primary measure of segregation; if the variance is zero there would be no segregation and each district would be the same as the average. We can thus compare the eight variances within and between levels to assess the degree of segregation and put credible intervals around these estimates to show the degree of support for different values of the estimates. Also in the random part of the model are a large number of covariances between the differentials at a particular scale. Thus σ_{v18} gives how British Salariat and Semi-/Unskilled co-vary at the District level. These can be converted into correlations by standardizing the covariance through dividing by the product of the square root of the associated variances. A positive value suggests that the two groups co-locate; a negative value suggests the two groups 'repel' each other. These correlations are accompanied by credible intervals so that we can statistically judge the degree of support for the estimated values. The model can be readily extended to another level for Region by including further differentials from the mean and summarizing these differences as variances and covariances associated with different types of people.

The key characteristic of a Poisson distribution is that the mean and variance are always exactly the same. This is achieved in the final line of the model specification by stating that the variance of the observed counts conditional on the underlying rate is equal to the underlying rate. Consequently, the other variance terms are estimated net of Poisson stochastic variation. In practice in this three-level model there is exactly the same set of units – known as the 'cells' – at level 1 and level 2; that is, each level 2 unit has exactly one level 1 unit. This views the aggregated counts at level 2 as consisting of replicated responses for 'individuals' at level 1. This use of a pseudo-level is explained in Browne et al. (2000) in relation to the binomial model and allows the separation of the variance into an exact Poisson at level 1 and over-dispersion at level 2 and level 3 so that the higher-level variances summarize the 'true' differences between areas over and above those expected from stochastic variation.[3]

This most complex model can readily be simplified by keeping the same observed counts but by reducing the number of parameters. Thus, the eight fixed effects represent two ancestries by four classes and we estimate this model on the basis that each group has its own mean and variance, the latter representing the degree of segregation. This multiplicative (class by ancestry) model can be simplified to an additive model (class plus ancestry):

$$O_{ijk} \sim Poisson(\pi_{ijk})$$

$$Log_e(\pi_{ijk}) = Log_e(E_{ijk}) + \beta_{1jk}Sal_{ijk} + \beta_{2jk}RNM_{ijk} + \beta_{3jk}Skill_{ijk} + \beta_{4jk}Unskill_{ijk} + \beta_{5jk}Aust_{ijk}$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1jk}$$
$$\vdots$$
$$\beta_{5jk} = \beta_5 + v_{5k} + u_{5jk}$$

$$\begin{bmatrix} v_{1k} \\ \vdots \\ v_{5k} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{v1}^2 & & \\ \vdots & \ddots & \\ \sigma_{v15} & \dots & \sigma_{v5}^2 \end{bmatrix})$$

$$\begin{bmatrix} u_{1jk} \\ \vdots \\ u_{5jk} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{u1}^2 & & \\ \vdots & \ddots & \\ \sigma_{v15} & \dots & \sigma_{u5}^2 \end{bmatrix})$$

$$Var(O_{ijk}|\pi_{ijk}) = \pi_{ijk}$$

In this model the between-District variance for the British Salariat is given directly by σ_{v1}^2 while σ_{v5}^2 is the variance for Australian ancestry at that level irrespective of class. To obtain the variance for Australian Salariat we will need to add these two variances together. This formulation represents a considerable simplification with many fewer parameters but still sees both aspects of ancestry and class as being important to degree of observed segregation.

This model can be further simplified by keeping the same observed counts but removing all terms associated with ancestry so that there are only underlying class differences for the four classes and residual Poisson-based stochastic variation. The segregation for each class at a level is given directly (e.g. σ_{v1}^2 is the between District segregation for the Salariat of both ancestries) with ancestry not seen as important.

$$O_{ijk} \sim Poisson(\pi_{ijk})$$
$$Log_e(\pi_{ijk}) = Log_e(E_{ijk}) + \beta_{1jk}Sal_{ijk} + \beta_{2jk}RNM_{ijk} + \beta_{3jk}Skill_{ijk} + \beta_{4jk}Unskill_{ijk} +$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1jk}$$
$$\vdots$$
$$\beta_{4jk} = \beta_4 + v_{4k} + u_{4jk}$$

$$\begin{bmatrix} v_{1k} \\ \vdots \\ v_{4k} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{v1}^2 & & \\ \vdots & \ddots & \\ \sigma_{v14} & \dots & \sigma_{v4}^2 \end{bmatrix})$$

$$\begin{bmatrix} u_{1jk} \\ \vdots \\ u_{4jk} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_{u1}^2 & & \\ \vdots & \ddots & \\ \sigma_{v14} & \dots & \sigma_{u4}^2 \end{bmatrix})$$

$$Var(O_{ijk}|\pi_{ijk}) = \pi_{ijk}$$

A further reduction in parameters is achieved in a model with no class differences in segregation and only those based on ancestry. In this model σ_{v1}^2 is the between-District variance for British ancestry and there are no terms for class whatsoever.

$$O_{ijk} \sim Poisson(\pi_{ijk})$$
$$Log_e(\pi_{ijk}) = Log_e(E_{ijk}) + \beta_{1jk}Brit_{ijk} + \beta_{2jk}Aust_{ijk}$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1jk}$$
$$\beta_{2jk} = \beta_2 + v_{2k} + u_{2jk}$$

$$\begin{bmatrix} v_{1k} \\ v_{2k} \end{bmatrix} \sim N(0 \begin{bmatrix} \sigma_{v1}^2 & \\ \sigma_{v12} & \sigma_{v2}^2 \end{bmatrix})$$

$$\begin{bmatrix} u_{1jk} \\ u_{2jk} \end{bmatrix} \sim N(0 \begin{bmatrix} \sigma_{u1}^2 & \\ \sigma_{u12} & \sigma_{u2}^2 \end{bmatrix})$$

$$Var(O_{ijk}|\pi_{ijk}) = \pi_{ijk}$$

Finally we arrive at the simplest possible null three-level model with just a constant (a set of 1s) representing everybody so that there is now a single overall mean (β_0) and a single variance at each level summarising differences around that mean net of stochastic variation. While this is not a substantively interesting model it provides a baseline for comparison with the other more complex models as the observed counts remains the same in all models and we can therefore judge changes in fit as complexity is introduced.

$$O_{ijk} \sim Poisson(\pi_{ijk})$$
$$Log_e(\pi_{ijk}) = Log_e(E_{ijk}) + \beta_{0jk}Constant_{ijk} +$$

$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$$

$$[v_{0k}] \sim N(0[\sigma^2_{v0}])$$

$$[u_{0jk}] \sim N(0[\sigma^2_{u0}])$$

$$Var(O_{ijk}|\pi_{ijk}) = \pi_{ijk}$$

All the models were estimated in MLwiN software as Fully Bayesian models using MCMC procedures (Browne, 2017;).[4] There are four reasons for using this approach in comparison to the more commonly applied maximum-likelihood procedure. First, we are likely to obtain less biased estimates and this is particularly important in discrete outcome models (Browne and Draper, 2006). The Full aspect (unlike Empirical Bayes) means that uncertainty in one parameter is taken into account in estimating all other parameters, thereby reducing potential bias. Second, we are particularly interested in the variance and correlation parameters and these can be anticipated to be non-normally distributed with the variance terms expected to be positively skewed as these estimates cannot go below zero; the correlations are probably either positively or negatively skewed as they approach -1 and +1 respectively. The MCMC procedure characterises these skewed distributions and we do not have to make asymptotic normality assumptions to characterise their uncertainty. Third, MCMC produces Bayesian credible and not frequentist confidence intervals which applies to the parameter and not the data, giving for example the 95% probability that the parameter falls between the lower and upper values. Moreover, as the distribution of parameter estimates can be non-normal it is possible to have asymmetric intervals that characterise the degree of empirical support for different values of the estimate. Finally, an important by-product of the estimation is the Deviance Information Criterion (DIC), a diagnostic for model comparison (Spiegelhalter et al., 2002). This is a badness-of-fit measure penalized for model complexity so its use is based on parsimony. Thus it is possible to fit all the model specifications outlined above and see to what extent the predicted counts based on estimated parameters come close to matching the observed counts. The degree of complexity is given by the degrees of freedom consumed in the model fit and this is estimated during the model-fitting process. Parameters like the mean and variances as usual are equivalent to one degree of freedom; however, the differentials at each level are also counted but may not individually contribute a whole degree of freedom as they come from a common distribution.

It is standard practice to regard a difference of 10 in the DIC as meaning the worse model has virtually no support and can be omitted from further consideration. Here we innovatively compare the improvement of fit by examining the changing DIC across a range of different models for segregation with same observed responses which are structured and produced according to the most complex model. The null multilevel model produces the baseline worst-fitting model and we can judge the extent to which different models with different effects for ancestry and class and (many) more parameters represent a genuine improvement over simpler models.

The distribution of the parameters in Bayesian modelling is known as the posterior and this characterises the degree of support for different values of the estimates. It is obtained by starting with a prior distribution, an initial guess at the distribution of the estimates, and then combining this with the likelihood distribution based on information contained in the data. The posterior distribution even in the simplest models is highly complex as it is the joint multivariate distribution of all the parameters (means, variances, covariances and differentials) considered simultaneously. Estimation works by making a simulated draw from a marginal distribution of one parameter and feeding this through into simulated draws for other parameters so that the full uncertainty in all parameters is taken into account. In practice we used priors with weak information so that the data have a greater effect on the posterior distribution. The means were given a uniform prior distribution and the variances were assumed to come from an inverse Wishart distribution, thereby allowing potential skewness in the parameter distribution. Maximal quasi-likelihood estimates obtained by MLwiN were used to get initial estimates to start the simulation. This was followed in all models by a discarded burn-in of 5000 simulated draws for each parameter (to get away from the potentially biased estimates) followed by a further 100,000 draws for each parameter to characterise the posterior.[5] The trajectories of these draws were inspected to see that there was no trending (that is failure to converge to the equilibrium posterior distribution) and that the effective sample size of each set posterior estimates was at least equivalent to 750 independent draws. The 2.5 and 97.5 percentiles of the posterior distributions were used to get the 95% credible intervals while the mean was used for the point estimates. The 100,000 estimates of the covariances and variances were stored and these were subsequently used to characterise the posterior distribution of the correlations.

The estimated variances on the log scale are transformed for ease of interpretation to Median Rate Ratios (MRRs):

$$MRR = \exp[\sqrt{2 * Variance} \; * 0.6745]$$

the value 0.6745 is the 75th percentile of the cumulative distribution function of the Normal distribution with mean 0 and variance 1. The credible intervals for a MRR are obtained by plugging in the credible intervals of the variance on the log scale obtained from the MCMC run.

Although change in the DIC is an absolute measure of the badness of fit, it is also possible to identify the relative importance of different models through the percentage change in the DIC. Table 3 shows the relative percentage change in comparison to the null model as the four more complicated models are fitted to each of the sixteen ancestry groups. There are clear differences between the ancestry groups in the relative importance of the various models. For six of the groups (British, New Zealand, Dutch, German, French and Polish) the percentage change for the first model is less than 10, indicating that ancestry alone does not account for a substantial part of their segregation; for five other groups (Russian, Yugoslav, Lebanese, Indian and Chinese), on the other hand, a reduction of more than 50 per cent indicates that ancestry alone is a very significant contributor to their residential segregation. The results from the second model show very large reductions in DIC for that first set of ancestry groups when occupational class is the only aspect considered in their segregation, but substantially smaller reductions for the latter set.

To summarise the patterns in Table 3 the sixteen ancestry groups were classified, using a hierarchical grouping algorithm, according to their percentage change in DIC profiles, which resulted in three separate groups. Their profiles across the five models are shown in Figure 2.

- The first group (Figure 2a) contains nine ancestries, comprising all of the English-speaking groups plus the three from western Europe, those with Polish ancestry and two of the three southern European groups. For them, model 1 – ancestry only – provides only a small reduction in DIC whereas model 2 – occupational class only – provides a very substantial decline and models 3 and 4, which combine the two variables, lead to only a slight further reduction in DIC. Within Sydney, relative to the residential distribution of the Australian host population, therefore, these nine ancestral groups are segregated by class but not also by ancestry: their residential distribution predominantly reflects their occupational patterns only.
- The second group (Figure 2b) contains five ancestry groups (Chinese, Indian, Lebanese, Russian and Yugoslav) for all of which model 1 – ancestry only – contributes to a much greater reduction in DIC than does model 2 – class only – with models 3 and 4 adding substantially to the reduction in DIC. These groups are substantially segregated into different parts of Sydney by their ancestry – i.e. there is substantial ethnic segregation – and, within that, by class also.
- The final group (Figure 2c) shows that the Greek and Korean ancestral groups are equally segregated by both class and ancestry, both separately and in conjunction.

This final stage of the modelling, substantially extending earlier work, clearly divides the sixteen ancestral groups analysed into two very distinct patterns of residential segregation. The first – comprising mostly ancestry groups that have been established in Sydney for several decades at least – have a pattern of segregation by occupational class that is not substantially different from that of those who claim Australian ancestry; theirs is a classic pattern of class segregation with only small elements of ethnic segregation as well. The second pattern, by contrast, characterises ancestral groups, many of whose members have moved to Sydney in large numbers relatively recently and who are also culturally more distinct from their host society than groups displaying the first pattern. They are substantially segregated by both ancestry and class; ethnic segregation within which there is also class segregation.

The pattern of segregation: sharing space

To what extent do members of the different occupational classes within each ancestry group share the same regional, district and suburban spaces? Our expectation is that they would be clustered in the same regions but within them to occupy separate districts and, especially, suburbs. Whether this is the case can be evaluated by the correlations derived from the MRR calculations. For the regional scale, those correlations – which vary between -1.0 and +1.0 and can be interpreted in the same way as standard product moment correlations – evaluate the closeness of two distributions of the modelled observed:expected ratios. (Each correlation coefficient has associated CIs, but these are not reported here.) At the district scale, they evaluate the closeness of those modelled ratios, net of the correlations at the regional scale; and at the suburb scale, net of the correlations at both of the larger scales.

Those correlations are shown in Table 3 and Figure 3. The same general pattern occurs at all three scales for many of the ancestry groups – the closer two occupational classes are on the continuum from the salariat at one end to the semi- and unskilled at the other, the higher the correlation, and thus the greater the likelihood of finding their members clustered in the same regions, districts and suburbs. Thus, for example, at the regional scale the correlation between the salariat and the routine non-manual class for the Irish was 0.56, whereas between the salariat and the skilled it was 0.06 and between the salariat and the semi-/unskilled it was -0.54. Between the adjacent routine non-manual and skilled groups it was 0.50, and it was 0.51 between the latter group and the semi-/unskilled, but between the routine non-manual and the semi-/unskilled it was only 0.03. This pattern is clearly shown in Figure 3's three graphs by the double-V shape of the trend of correlations across the six comparisons

At the regional scale, the average correlation for pairs of adjacent occupations on the continuum was 0.58; for pairs separated by one class (e.g. between the salariat and the skilled) it was 0.39; and for pairs two classes apart (salariat and semi-/unskilled) it was 0.05. The comparable figures at the district scale were 0.50, 0.40 and 0.19; and at the smaller suburban scale they were 0.56, 0.47 and 0.27. Within each ancestry group, relatively similar occupational classes were more likely to be found clustered in the same regions, within those regions within the same districts, and within those districts within the same suburbs, than were dissimilar classes. There were exceptions, however, as picked out by the trends on the graphs that either lack the characteristic double-V shape or for which it is much less pronounced than it is for most groups. At the regional scale, six ancestral groups stand out as separate, with high and relatively invariant correlations between all pairs – Greek, Maltese, Lebanese, Indian, Chinese and Korean: all occupational classes within those ancestries tend to be concentrated in the same regions. At the district and suburb scales, only three ancestry groups – Indian, Chinese and Korean – stand out: their members are concentrated in the same smaller areas of the residential mosaic whatever their occupational class, whereas for most others concentration in the same districts and suburbs within regions is much less.

Clustering of the seventeen ancestry groups according to their correlation profiles at all three scales generated three clear groupings:

- Australian, British, Irish, New Zealand, German, French, Italian;
- Polish, Russian, Yugoslav; and
- Greek, Maltese, Lebanese, Indian, Chinese, Korean.

Their average profiles across the three scales for the six correlations are in Figure 4, which shows considerable variation in the intensity of the double-V shape. The differences are greatest for the first group – comprising the English-speaking and (mostly) north European ancestries – and least for the third – comprising southern and eastern Mediterranean and the three Asian ancestries. For the former group, there is substantial occupational class segregation; for the latter, relatively little.

Conclusions

This paper has used data on the occupational structure of seventeen ancestry groups in the Sydney urban region to illustrate the application of a major extension to a recently-developed statistical modelling procedure to studies of the joint influence of ancestry (as a surrogate for ethnicity) and occupational class on patterns of residential segregation there, at three nested spatial scales. The results show that for some ancestral groups – mainly those longest-established in Sydney and with few major cultural differences from the host society – occupational class is the main influence on their residential location. For other groups – predominantly comprising recent migrants to Australia and their dependents, who differ more from the local cultural norms – ancestry is the main influence and there is relatively little segregation by class within each ancestral group. Additionally, among the latter there is more sharing of the same spaces by members of different occupational classes, at all scales, than is the case with the former. Class and ethnicity play different roles in the structuring of Sydney's residential mosaic depending on the ancestral groups concerned; there are some general patterns reflecting both the average length of time that group members have been resident in Australia and their cultural characteristics (the three Asian groups have different patterns from those from western Europe, for example), but also some group-specific influences. These findings are, of course, specific to Sydney, but its labour and housing markets are typical of those in most contemporary, multi-ethnic and multi-cultural world cities; they set a benchmark against which the situation in other cities can be compared.

It is widely recognised that residential segregation is a multi-faceted feature of contemporary cities, but means of establishing the relative importance of various influences on that segregation, at various spatial scales, have not been available. The method outlined, and substantially extended, here provides a means of filling that gap, and is potentially widely applicable in investigations of hierarchically-structured spatial patterns. Using it the possibility that segregation patterns may be wrongly attributed to the wrong influences – class rather than ethnicity, for example, or religion rather than education – is reduced, increasing the potential of appreciating the underlying processes leading to residential patterns from ecological data.

Notes

---

[1] We use the TableBuilder Pro version of TableBuilder, on which see http://www.abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder – accessed 4 April 2017. Data from TableBuilder Pro downloads are licensed and the product is charged for.

[2] There would be 17 * 4, that is 68, means of the (log) of observed:expected ratios, 68 variances at each of the three levels, 2,278 covariances ((68^2 -68) * ½) at each of the three levels, the 68 * 43 modelled rates at the macro level, the 68 * 223 modelled rates at the meso level and the 68* 616 modelled rates ate the micro level.

[3] As Lee et al. (2015) have shown, it is important to take spatial autocorrelation into account when modelling segregation, both substantively and for correct measures of uncertainty. Currently, the model achieves this in a rather crude way in that suburbs are nested in districts and this implicitly models dependency, as shown in Jones et al (2015). A submitted paper (Jones et al. 2017) show how explicit spatial autocorrelation can be accommodated in the model through a multilevel multiple-membership structure where each suburb has its own surrounding bespoke neighbourhood.

[4] Pillinger(undated) provides a detailed account of how MCMC estimation in MLwiN can be used for modelling segregation when the outcome is a proportion while Jones and Subramanian (2013) detail the use of the software for the Poisson multilevel modelling of counts.

[5] Draper (2008) in his good practice recommendations suggests a burn-in of 500 followed by 5000 monitoring draws but we have found that models using the Poisson distribution to be quite correlated so the posterior distribution is explored quite slowly. We have therefore increased both burn-in and monitoring substantially.

References

Browne W J (2017) *MCMC Estimation in MLwiN 3.0,* Centre for Multilevel Modelling, University of Bristol. Available at http://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/3-00/mcmc-web.pdf; last accessed online 25/03/2017

Browne W J and Draper D (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis* 1: 473-550.

Browne W J, Subramaniam S V, Jones K and Goldstein H (2005) Variance partitioning in multilevel logistic models that exhibit overdispersion. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168: 99-113.

Butler T and Hamnett C (2011) *Ethnicity, Class and Aspiration: Understanding London's New East End*. Bristol: Polity Press.

Darroch A G and Marston W G (1971) The social class basis of ethnic residential segregation: the Canadian case. *American Journal of Sociology* 77: 491-501.

Draper D (2008) Bayesian multilevel analysis and MCMC. In Leeuw, J and Meijer, E (eds.), *Handbook of Multilevel Analysis*, Springer.

Duncan C and Jones K (2001). Using multilevel models to model heterogeneity: potential and pitfalls. *Geographical Analysis* 32: 279-305.

Iceland J, Sharpe C and Steinmetz E (2005) Class differences in African-American residential patterns in US metropolitan areas. *Social Science Research* 34: 252-266.

Intrator J, Tannen J and Massey DS (2016) Segregation by race and income in the United States 1970-2010. *Social Science Research* 40: 45-60.

Johnston R J, Forrest J, Jones K. and Manley D (2016) The scale of segregation: ancestral groups in Sydney, 2011. *Urban Geography*, doi 10.1080/02723638.2016.1139867

Johnston R J, Forrest J, Manley D and Jones K (2017) The segregation of generations: ancestral groups in Sydney 2011. *Geographical Research*, doi 10.1111/1745-5871.12233

Jones K, Johnston R J, Manley D, Owen D and Charlton C (2015) Ethnic residential segregation: a multi-level, multi-group, multi-scale approach – exemplified by London in 2011. *Demography* 52: 1995-2019.

Jones K, Manley D and Johnston R J (2017) Modelling residential segregation as unevenness and clustering: a multilevel approach incorporating spatial dependence and tackling the MAUP. Submitted (copies available from authors).

Jones K and Subramanian S V (2013) *Developing multilevel models for analysing contextuality, heterogeneity and change using MLwiN Volume 2*. Bristol: Centre for Multilevel Modelling, University of Bristol. Available at https://www.researchgate.net/publication/

260772180_Developing_multilevel_models_for_analysing_contextuality_heterogeneity_and_change_using_MLwiN_Volume_2, last accessed online 26/7/2017.

Kapoor N (2013) Rethinking empirical approaches to racial segregation. *Sociological Review* 61: 440-459.

Lee D, Minton J and Pryce G (2015) Bayesian inference for the Dissimilarity Index in the presence of spatial autocorrelation. *Spatial Statistics* 11: 81-95.

McCullagh P and Nelder J A (1989) *Generalized Linear Models*. London: Chapman & Hall.

Manley D, Johnston R J, Jones K and Owen D (2015) Macro-, meso- and microscale segregation: modelling changing ethnic residential patterns in Auckland, New Zealand, 2001-2013. *Annals of the Association of American Geographers* 105: 951-967.

Massey D S (1981) Social class and ethnic segregation: a reconsideration of methods and conclusions. *American Sociological Review* 46: 641-650.

Massey D S, Rothwell J and Domina T (2009) The changing bases of segregation in the United States. *Annals of the American Academy of Political and Social Sciences* 626: 74-90.

Pillinger R (undated) *Modelling ethnic segregation using MLwiN*. Available at http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/modelling-segregation-mlwin.pdf; last accessed online 26/7/2017.

Reardon S F, Townsend J and Fox L (2017) A continuous measure of the joint distribution of race and income among neighbourhoods. *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3 (2): 34-62.

Spiegelhalter D J, Best N G, Carlin B P and van der Linde A (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 64: 583-640.

Table 1. The occupational structure of the ancestry groups analysed

	Salariat	Routine Non-Manual	Skilled	Semi-/ Unskilled
Australian	173,161	146,336	50,101	44,235
British	299,763	211.568	69.351	58,943
Irish	87,647	57,785	17,698	14,173
New Zealand	7,457	5,372	1,810	1,693
Dutch	10,696	7,014	2,629	1,789
German	25,789	17,572	6,029	4,633
French	6,724	4,208	1,174	911
Polish	8,657	5,888	2,344	1,761
Russian	7,230	3,879	659	614
Yugoslav	10,411	26,260	5,730	5,980
Italian	31,697	29,764	12,213	8,803
Greek	19,098	16,353	6,322	5,294
Maltese	7,490	8,520	3,881	4,318
Lebanese	15,045	14,811	8,219	7,151
Indian	32,756	21,551	6,389	9,473
Chinese	71,008	49,838	17,214	21,814
Korean	7,823	5,197	3,358	3,131

Table 2. The MRR values for each occupational class in each ancestry group at each scale, with their associated CIs

Regional Scale

	Salariat			RNM			Skilled			Semi-Unskilled		
Ancestry	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	LCI
Australian	1.36	1.47	1.62	1.27	1.35	1.45	1.53	1.70	1.94	1.99	2.38	2.94
British	1.55	1.74	2.00	1.23	1.30	1.39	1.38	1.50	1.66	1.74	2.00	2.37
Irish	1.61	1.82	2.11	1.19	1.25	1.33	1.27	1.37	1.49	1.57	1.78	2.07
New Zealand	2.27	2.87	3.82	1.73	2.15	2.77	1.64	2.26	3.26	1.91	2.76	4.19
Dutch	1.82	2.14	2.61	1.68	2.01	2.48	2.50	3.43	4.98	5.31	9.54	19.34
German	1.61	1.82	2.11	1.22	1.30	1.40	1.37	1.52	1.73	2.29	2.93	3.93
French	2.62	3.47	4.85	1.77	2.27	3.04	2.28	3.70	6.58	1.50	2.53	5.26
Polish	1.64	1.90	2.28	1.43	1.69	2.05	2.07	2.71	3.73	3.19	5.35	9.69
Russian	2.16	2.75	3.68	1.82	2.26	2.93	2.15	3.10	4.79	2.62	4.36	8.03
Yugoslav	1.26	1.46	1.77	1.27	1.46	1.76	1.32	1.53	1.87	1.33	1.59	2.03
Italian	1.47	1.64	1.88	1.55	1.75	2.03	1.73	2.02	2.43	2.37	3.02	4.03
Greek	2.36	3.00	4.01	2.84	3.78	5.33	3.62	5.23	8.17	5.45	9.01	16.62
Maltese	1.70	2.04	2.56	2.63	3.49	4.89	4.42	6.87	11.63	10.04	19.99	45.58
Lebanese	3.10	4.27	6.33	3.74	5.37	8.36	5.18	8.24	14.40	8.12	15.26	32.39
Indian	2.08	2.59	3.00	2.27	2.91	3.89	2.61	3.54	5.07	3.57	5.32	8.56
Chinese	2.53	3.29	4.49	2.64	3.54	5.00	2.92	4.07	5.99	3.89	5.90	9.89
Korean	6.01	10.47	20.52	8.59	17.97	43.26	8.54	19.03	49.31	10.50	26.12	78.72

District Scale

	Salariat			RNM			Skilled			Semi-Unskilled		
Ancestry	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	LCI
Australian	1.18	1.22	1.26	1.14	1.17	1.20	1.17	1.22	1.27	1.25	1.32	1.39
British	1.21	1.25	1.29	1.13	1.16	1.18	1.16	1.20	1.25	1.20	1.25	1.31
Irish	1.22	1.27	1.32	1.14	1.18	1.21	1.16	1.21	1.27	1.24	1.33	1.43
New Zealand	1.43	1.63	1.87	1.27	1.43	1.66	1.28	1.50	1.82	1.29	1.56	2.02
Dutch	1.14	1.24	1.36	1.28	1.50	1.75	1.30	1.61	2.06	1.27	1.55	2.06
German	1.18	1.25	1.31	1.08	1.14	1.20	1.20	1.31	2.43	1.21	1.36	1.56
French	1.34	1.58	1.88	1.09	1.19	1.37	1.09	1.22	1.50	1.37	2.02	4.37
Polish	1.26	1.38	1.52	1.27	1.46	1.72	1.27	1.47	1.76	2.08	3.15	5.27
Russian	1.43	.161	1.87	1.47	1.69	2.04	1.55	2.10	3.10	2.30	3.20	4.71
Yugoslav	1.24	1.39	1.63	1.24	1.38	1.59	1.38	1.56	1.82	1.31	1.53	1.88
Italian	1.39	1.47	1.52	1.36	1.43	1.52	1.37	1.47	1.58	1.60	1.81	2.06
Greek	1.64	1.82	2.03	1.59	1.79	20.4	1.57	1.83	2.17	1.70	2.13	2.80
Maltese	1.80	2.07	2.41	1.73	1.99	2.31	1.79	2.17	2.64	2.25	2.91	3.91
Lebanese	1.66	1.86	2.09	1.98	2.33	2.75	2.37	2.88	3.56	3.19	4.25	5.82
Indian	1.73	1.93	2.16	1.89	2.12	2.40	2.20	2.55	2.99	2.71	3.26	4.00
Chinese	1.64	1.78	1.94	1.80	1.98	2.20	1.91	2.15	2.44	2.25	2.63	3.09
Korean	2.23	2.76	3.50	2.83	3.76	5.35	3.41	5.13	7.88	3.52	5.19	8.07

Suburb Scale

Ancestry	Salariat			RNM			Skilled			Semi-Unskilled		
	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	HCI	LCI	MRR	LCI
Australian	1.20	1.22	1.25	1.19	1.20	1.22	1.26	1.29	1.32	1.43	1.48	1.54
British	1.24	1.26	1.28	1.17	1.19	1.21	1.27	1.30	1.33	1.35	1.39	1.43
Irish	1.26	1.28	1.32	1.20	1.23	1.26	1.32	1.37	1.42	1.58	1.66	1.75
New Zealand	2.18	2.45	2.77	3.43	4.03	4.80	6.48	8.63	11.82	15.48	25.16	44.15
Dutch	1.87	2.03	2.22	2.74	3.13	3.60	4.16	5.18	6.62	11.50	17.89	28.68
German	1.30	1.35	1.40	1.35	1.41	1.47	1.85	2.01	2.19	3.34	3.91	4.61
French	2.71	3.13	3.63	4.79	5.87	7.35	12.49	19.71	33.50	37.94	87.99	236.78
Polish	1.97	2.15	2.36	3.31	3.85	4.51	4.59	5.68	7.16	12.96	21.39	38.99
Russian	3.38	3.79	4.26	4.75	5.59	6.65	10.95	17.05	27.70	14.59	25.39	45.43
Yugoslav	7.11	8.38	9.97	7.01	8.15	9.57	9.51	11.66	14.54	15.55	20.51	27.69
Italian	1.44	1.49	1.55	1.50	1.56	1.63	1.81	1.92	2.05	2.50	2.80	3.15
Greek	1.94	2.10	2.29	2.46	2.73	3.06	2.73	3.17	3.69	4.06	5.00	6.26
Maltese	2.53	2.81	3.15	2.56	2.88	3.27	3.77	4.51	5.47	4.34	5.48	7.08
Lebanese	2.31	2.55	2.84	2.86	3.27	3.78	3.12	3.69	4.39	4.45	5.69	7.84
Indian	1.95	2.08	2.23	2.08	2.24	2.42	2.39	2.67	3.01	3.11	3.62	4.24
Chinese	1.70	1.79	1.90	1.92	2.05	2.19	2.17	2.38	2.63	2.75	3.09	3.51
Korean	3.14	3.74	4.53	4.19	5.50	7.42	5.13	7.30	10.68	5.44	7.72	11.26

Table 3. The percentage change in the DIC relative to the null model for each ancestry group's segregation

	Null	Ancestry Only	Class Only	Class + Ancestry	Class * Ancestry
British	100	94	27	23	12
Irish	100	88	41	32	31
New Zealand	100	94	42	37	30
Dutch	100	97	41	38	32
German	100	96	40	37	34
French	100	92	43	36	28
Polish	100	93	43	36	30
Russian	100	31	73	4	4
Yugoslav	100	47	72	21	17
Italian	100	71	57	28	26
Greek	100	61	63	24	21
Maltese	100	81	49	32	27
Lebanese	100	48	69	20	17
Indian	100	49	69	19	17
Chinese	100	37	76	14	12
Korean	100	62	61	23	20

Table 4. The correlations between the distributions between each pair of occupational classes within each ancestry group at each scale. (Key to occupations: Sal – Salariat; RNM- Routine Non-Manual: Skill – Skilled; SSU – Semi-/Unskilled.)

Regional Scale

Ancestry	Sal/RNM	Sal/Skill	Sal/SSU	RNM/Skill	RNM/SSU	Skill/SSU
Australian	0.04	-0.17	-0.48	0.76	0.64	0.76
British	0.34	0.03	-0.56	0.65	0.74	0.65
Irish	0.56	0.0.6	-0.54	0.55	0.03	0.51
New Zealand	0.74	0.35	-0.57	0.50	-0.34	0.08
Dutch	0.60	0.36	-0.12	0.55	0.12	0.39
German	0.31	-0.26	-0.57	0.49	0.24	0.64
French	0.63	0.49	0.06	0.54	0.35	0.27
Polish	0.35	-0.29	-0.28	0.30	0.34	0.76
Russian	0.69	0.52	-0.14	0.66	0.13	0.26
Yugoslav	0.67	0.42	-0.28	0.52	0.38	0.40
Italian	0.46	0.18	-0.02	0.59	0.52	0.58
Greek	0.68	0.49	0.46	0.69	0.72	0.76
Maltese	0.59	0.59	0.56	0.87	0.87	0.90
Lebanese	0.73	0.65	0.64	0.78	0.78	0.84
Indian	0.75	0.65	0.58	0.79	0.77	0.79
Chinese	0.87	0.78	0.72	0.86	0.83	0.86
Korean	0.89	0.84	0.84	0.87	0.86	0.87

District Scale

Ancestry	Sal/RNM	Sal/Skill	Sal/SSU	RNM/Skill	RNM/SSU	Skill/SSU
Australian	0.59	0.36	-0.18	0.77	0.45	0.60
British	0.67	0.42	-0.31	0.73	0.24	0.47
Irish	0.70	0.43	-0.10	0.70	0.33	0.57
New Zealand	0.22	0.06	-0.14	0.03	0.05	0.08
Dutch	0.56	0.48	0.11	0.42	0.19	0.12
German	0.61	0.57	0.01	0.52	0.22	0.29
French	0.07	0.28	0.08	0.06	0.09	0.06
Polish	0.44	0.17	0.04	0.18	-0.04	0.22
Russian	0.44	0.32	0.09	0.46	0.24	0.22
Yugoslav	0.67	0.33	0.29	0.49	0.38	0.37
Italian	0.71	0.64	0.32	0.64	0.50	0.43
Greek	0.55	0.28	0.26	0.42	0.40	0.47
Maltese	0.58	0.44	0.38	0.45	0.56	0.45
Lebanese	0.55	0.39	0.44	0.60	0.60	0.64
Indian	0.83	0.73	0.66	0.79	0.73	0.73
Chinese	0.77	0.61	0.65	0.78	0.76	0.77
Korean	0.65	0.67	0.68	0.66	0.67	0.67

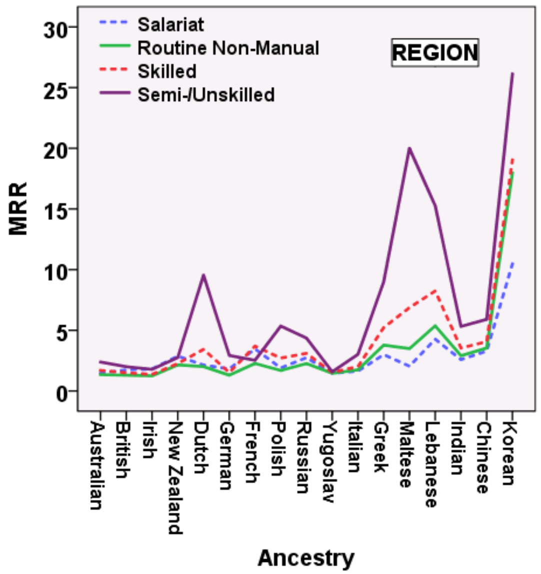
Suburb Scale

Ancestry	Sal/RNM	Sal/Skill	Sal/SSU	RNM/Skill	RNM/SSU	Skill/SSU
Australian	0.61	0.38	-0.03	0.85	0.60	0.74
British	0.61	0.31	-0.19	0.72	0.48	0.59
Irish	0.56	0.28	-0.19	0.67	0.32	0.59
New Zealand	0.20	0.18	0.11	0.31	0.31	0.31
Dutch	0.12	0.11	-0.09	0.12	0.09	0.20
German	-0.06	0.13	-0.14	0.09	0.16	0.18
French	0.01	0.02	-0.01	0.37	0.26	0.24
Polish	0.85	0.74	0.82	0.90	0.83	0.74
Russian	0.85	0.74	0.82	0.90	0.83	0.74

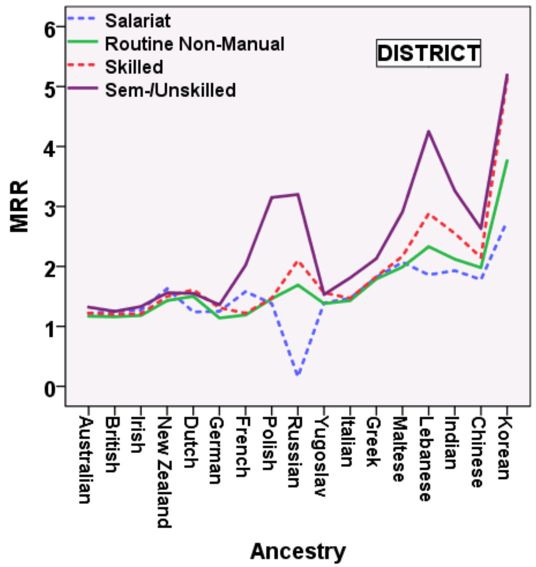
Yugoslav	0.96	0.86	0.75	0.93	0.83	0.89
Italian	0.57	0.45	0.09	0.66	0.54	0.50
Greek	0.36	0.28	0.19	0.59	0.43	0.53
Maltese	0.35	0.39	0.25	0.29	0.48	0.26
Lebanese	0.50	0.40	0.43	0.56	0.55	0.57
Indian	0.82	0.71	0.61	0.81	0.76	0.67
Chinese	0.73	0.64	0.55	0.81	0.77	0.82
Korean	0.57	0.59	0.67	0.66	0.66	0.60

Figure 1. The MRR values for each ancestry group and occupation at each scale: (a) Region; (b) District; (c) Suburb.

a)



b)



c)

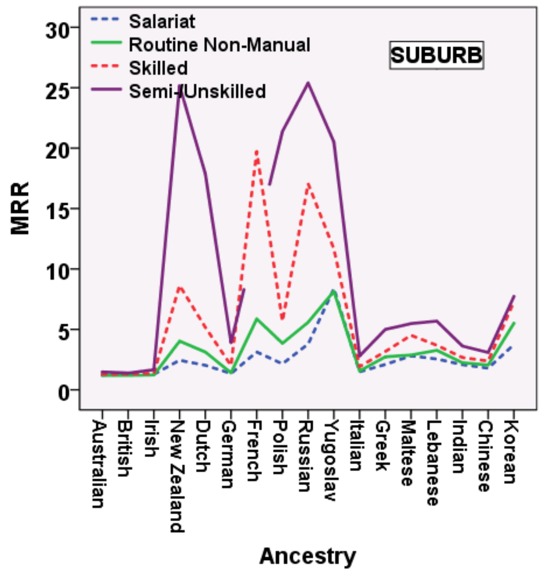


Figure 2. The profiles of each ancestry group in the three groups identified in their profiles in the reduction of the DIC.

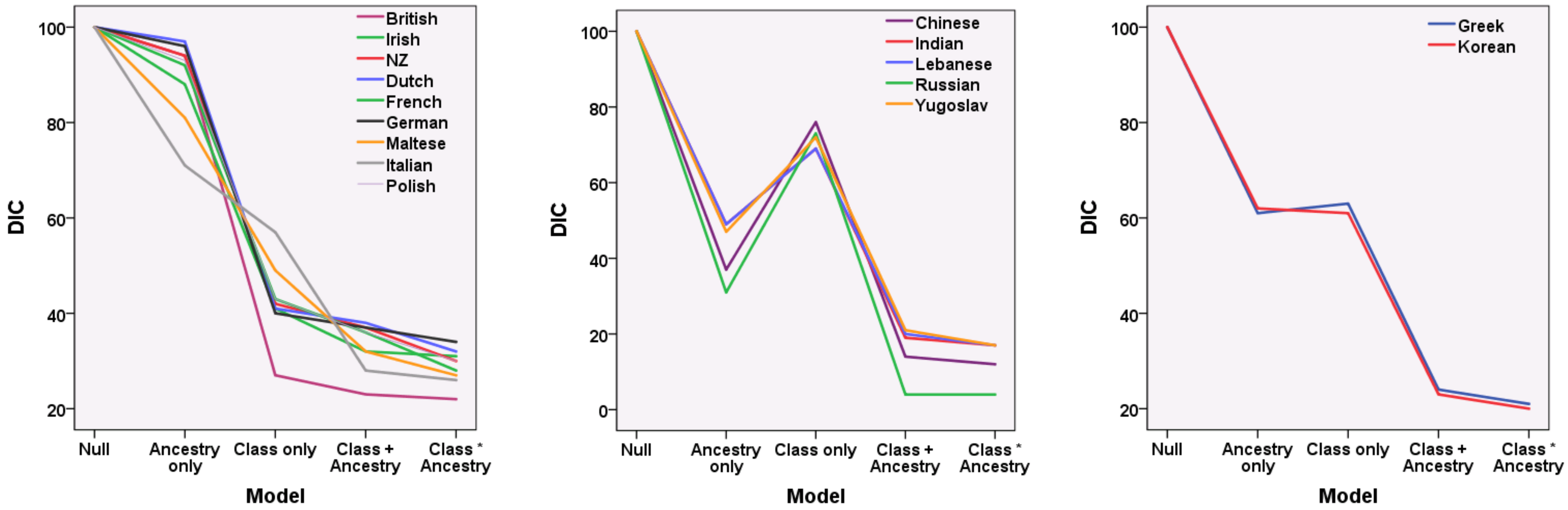
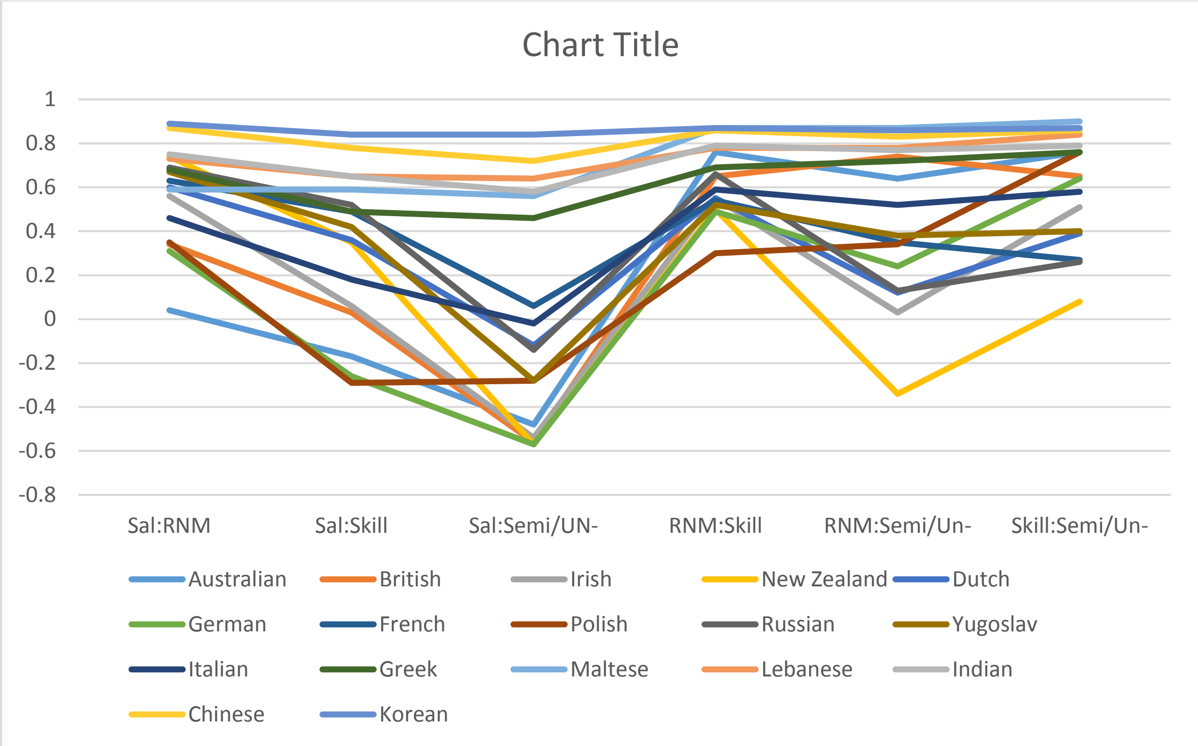
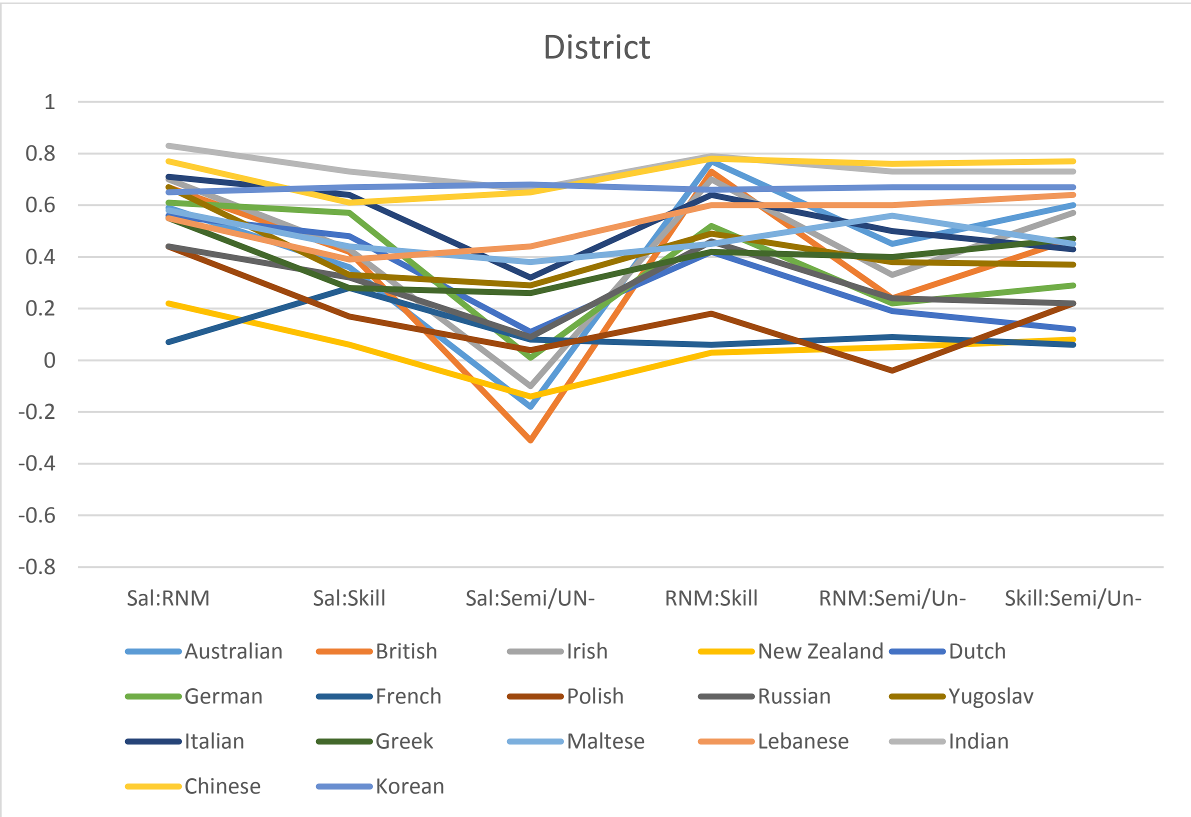
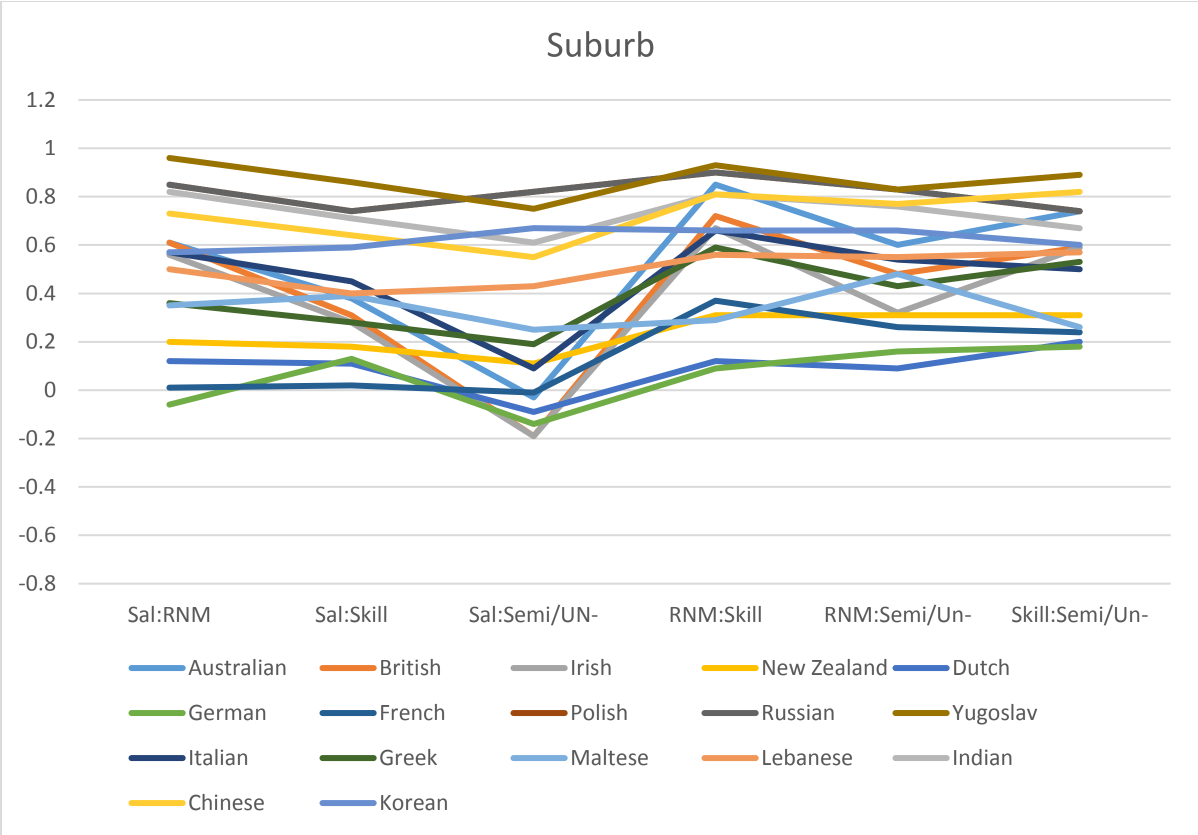


Figure 3. The correlations for pairs of occupations by ancestry group, at each scale: (a) region; (b) District; (c) Suburb. (Key to occupations: Sal – Salariat; RNM- Routine Non-Manual: Skill – Skilled; SSU – Semi-/Unskilled.)







Suburb
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Sal:RNM
Sal:Skill
Sal:Semi/UN-
RNM:Skill
RNM:Semi/Un-
Skill:Semi/Un-
Australian
British
Irish
New Zealand
Dutch
German
French
Polish
Russian
Yugoslav
Italian
Greek
Maltese
Lebanese
Indian
Chinese
Korean

Figure 4.The mean correlations for pairs of occupations for each group of ancestry groups at all scales combined. (Key to occupations: Sal – Salariat; RNM- Routine Non-Manual: Skill – Skilled; SSU – Semi-/Unskilled.)